

УДК 004.4

О.В. Лактіонов

Науковий керівник – Мелешко Є.В., канд. техн. наук, доцент
Кіровоградський національний технічний університет

Розробка програмного забезпечення системи комп’ютерного зору для створення інтерактивних користувальницьких інтерфейсів

У цей час завдання оцифрування й зберігання великих обсягів візуальної інформації мають закінчені технічні рішення, що цілком задовольняють вимогам користувачів, у той час як в області розробки методів рішення завдань пошуку й семантичної класифікації зображень ситуація виявляється прямо протилежною. Відзначимо, що до останнього часу найбільше часто використовувався пошук візуальної інформації, що опирається на індексування текстових описів, асоційованих із зображенням (наприклад, пошук картинок на Яндекс і Google). При очевидній необхідності організації доступу до колекції зображень за допомогою пошуку за текстовою інформацією, асоційованою із зображеннями, даний підхід представляється недостатнім. Дійсно, існуюча неоднозначність при встановленні відповідності між візуальним змістом і текстовим описом знижує показники точності й повноти пошуку, а цілому ряді випадків виявляється досить важким або взагалі неможливим скласти словесний опис зображення (наприклад, абстрактні картини).

На початку 80-х років для подолання недоліків пошукових систем на основі тексту були початі розробки методів пошуку зображень за змістом (у закордонній літературі для позначення даного підходу використовується абревіатура CBIR – Content-based image retrieval). В CBIR-системах зображення індексуються по їхньому візуальному вмісті (за кольором, текстурою, формою й т.д.).

Аналіз сучасних тенденцій розвитку систем пошуку зображень показує, що найбільш популярним виявляється підхід, заснований на використанні тих або інших алгоритмів аналізу зображень. Наприклад, одним з останніх нововведень Google стало введення нових типів зображень, які Google автоматично класифікує. На сучасний момент дана пошукова система розпізнає наступні типи зображень: портрети, фотографії, гравюри, олівцеві рисунки й схеми, «кліпарт». У квітні 2009 року запущений пошук схожих зображень Google Similar Images: на першому етапі використовується традиційний спосіб пошуку по текстовому запиті, а потім як запит використовується одне зі знайдених зображень і здійснюється пошук по візуальній подібності. У пошуковій системі Яндекс наприкінці 2008 року став можливий пошук портретів, а раніше був запущений пошук картинок по переважному кольору, пошук фотографій і механізм виявлення дублікатів.

У цей час відомі демо-версії систем пошуку зображень за змістом (MFIRS, CIREs, Tiltomo, INRIA, Retrievr і ін.). Проведений аналіз якості пошуку в перерахованих системах показує, що якість пошуку в них істотно нижче в порівнянні із системами пошуку зображень по текстових анотаціях. Так, з 20-ти перших знайдених зображень запиту відповідає не більше 16% зображень (при пошуку по текстових анотаціях у системі Яндекс аналогічний показник рівнявся 51%).

Слід зазначити, що завдання пошуку зображень за змістом по своїй постановці формально близька до завдання розпізнавання образів, однак по своїй суті ці завдання не є ідентичними. У завданні розпізнавання образів основною метою є віднесення вхідного зображення до одному із заздалегідь відомих класів, у той час як у завданні пошуку зображень за змістом споконвічно явної вимоги до ідентифікації класу вхідного зображення не ставиться, але потрібно знайти зображення, що володіють візуальною подібністю із запитом.

Аналіз базових принципів, використовуваних у сучасних CBIR-системах, показує, що в більшості з розглянутих систем пошуку зображень за змістом пошук ведеться за ознаками, що витягається із усього зображення (глобальні ознаки). Крім того, при пошуку зображень користувач найчастіше цікавиться цілком конкретним об'єктом, що є присутнім на зображенні, і, відповідно, результатом пошуку повинні бути зображення, що містять шуканий об'єкт.

У той же час у науковій літературі описані результати дослідження механізмів сприйняття людиною зображень, згідно яким увага людини в процесі аналізу зображень концентрується не на всьому зображенні, а на деяких цілком конкретних областях зображення. При цьому вибір області (або декількох областей), що привертає увагу людини, більшою мірою обумовлений властивостями даної області, але в меншому ступені високорівневими когнітивними процесами (модель висхідної уваги).

Мета й завдання дослідження. Ціль роботи складається в розробці програмного забезпечення, яке призначено для системи комп'ютерного зору для створення інтерактивних користувацьких інтерфейсів.

Для досягнення поставленої мети визначена програма дослідження, що складається з наступних завдань:

1. Аналіз сучасного стану проблеми пошуку зображень за змістом й огляд інформаційних ознак зображень, використовуваних в CBIR-системах.
2. Аналіз критеріїв якості сегментації й виділення критеріїв, які доцільно використовувати для оцінки якості алгоритмів сегментації зображень.
3. Аналіз відомих моделей висхідної візуальної уваги й дослідження ступеня суб'єктивності уваги людини.
4. Розробка алгоритму автоматичного знаходження об'єкту на зображенні, що привертає увагу людини.
5. Розробка методів пошуку зображень із використанням інформації про об'єкт і проведення експериментальної перевірки їхньої працездатності.
6. Розробка на основі запропонованих методів пошуку зображень прототипу CBIR-системи.

Об'єктом дослідження є процес створення системи комп'ютерного зору.

Предметом дослідження є системи комп'ютерного зору для створення інтерактивних користувацьких інтерфейсів.

Список літератури

1. Фершильд М.Д. Модели цветового восприятия, Рочестерский технологический институт, Манселловская научная лаборатория по цвету, 2006.
2. Pass G., Zabih R. Histogram refinement for content-based image retrieval // IEEE Workshop on Applications of Computer Vision, 1996, pp. 96-102.
3. Ma W.Y., Zhang H. Benchmarking of image feature for content-based retrieval // In IEEE 32nd Asilomar Conference on Signals, Systems, Computers, 1998, v. 1, pp. 253-257.
4. Deng Y., Manjunath B. S., Kenney Ch., Moore M. S., Shin H. An efficient color representation for image retrieval // IEEE Transactions on image processing, 2001, v. 10, no. 1, pp. 140-147.
5. Стандарт MPEG-7 [Електронний ресурс]: Дескриптор домінантного цвета. – Режим доступа: http://book.itep.ru/2/25/mpeg_7.htm, 27.03.2009.